

## VOWEL LENGTH IN INFANT-DIRECTED SPEECH: THE REALISATION OF SHORT-LONG CONTRASTS IN CZECH IDS

Michaela Svoboda<sup>a,b</sup>, Kateřina Chládková<sup>a,b</sup>, Tanja Kocjančič Antolík<sup>b,c,d</sup>, Nikola Paillereau<sup>b,c</sup>, Petra Slížková<sup>b</sup>

<sup>a</sup>Institute of Czech Language and Theory of Communication, Faculty of Arts, Charles University; <sup>b</sup>Institute of Psychology, Czech Academy of Sciences; <sup>c</sup>Institute of Phonetics, Charles University; <sup>d</sup>Faculty of Education, University of Ljubljana

michaela.svoboda@post.cz

### ABSTRACT

When interacting with young children, talkers across many languages use a speech style that reflects positive affect, draws infants' attention, and supposedly facilitates language acquisition. As for the latter, a well-documented feature of infant-directed speech is an exaggeration of spectrally-cued vowel contrasts. Here we tested whether talkers exaggerate also durationally cued contrasts.

Sixty-three mothers, native speakers of Czech, were recorded while playing with their infant (4- to 10-month-olds, IDS) and while speaking to an adult (ADS). The durations of the five Czech phonemically short vowels were compared to their long counterparts. Vowel duration (normalised for word duration) was longer in IDS than in ADS more for phonemically long vowels at the younger infant ages, indicating a developmentally specific early exaggeration of length contrasts in Czech infant-directed speech. The present finding suggests that in a language with phonemic length, caregivers' realisation of speech sounds may go beyond merely being longer and slower overall.

**Keywords:** infant-directed speech, vowel length, development of early input, Czech

### 1. INTRODUCTION

This paper investigates the realisation of phonemic short-long contrasts in Czech-speaking mother's infant-directed speech (IDS). Across many cultures and languages, adults use a distinct speech style when communicating with infants and young children. Recently conducted a meta-analysis reflecting results of 87 unique IDS studies on more than 20 different languages or language varieties [1] demonstrated that IDS often differs from adult-directed speech (ADS) in several aspects, including acoustic parameters such as mean F0 (which tends to be higher in IDS than in ADS), F0 range or standard deviation (larger in IDS than in ADS), speech tempo (slower in IDS),

the acoustic duration of vowels (longer in IDS), and the size of the vowel space (larger in IDS). Some of these features occur already in the speech addressed to unborn infants [2]. Although, in general, vowel spaces tend to be exaggerated in IDS compared to ADS, the literature also reports different tendencies for some languages. For instance, [2] did not find enlarged vowel spaces in Dutch mothers' IDS but reported vowel fronting which she attributed to the mothers' positive affect, i.e. smiling, while speaking to their child.

Researchers have argued that the exaggeration of the vowel space facilitates young children's acquisition of native phoneme contrasts (*the hyperarticulation hypothesis*, see [4, 5]). Whereas the literature provides much data on vowel spectral properties and enlargement (or absence or even shrinkage [3, 7]) of vowel quality contrasts [1], less is known about potential exaggeration of length contrasts. If acoustic enhancement occurs to aid children in learning native contrasts, one would expect to find the exaggeration also for vowel length distinctions in languages whose phonology contains phonemic length. However, data from languages with phonemic length are mixed. Some studies report exaggerated duration differences between short and long vowels in IDS (e.g. [6] for Swedish) while others only report longer duration overall indexing the slower speech tempo typical of IDS [e.g. 7 for Norwegian]. In Japanese IDS, an exaggerated length difference was found to be context-specific, reportedly occurring only in word-final positions [8]. Interestingly, Japanese infants seem to acquire context-independent native length contrasts relatively late (at about 9.5 months of age [9]), which one could potentially attribute to insufficient exposure to the long member of length contrasts in the input [10] or even to the non-exaggerated nature of short-long distinctions (outside the word-final position).

Here we assess the realisation of short-long vowel contrasts in Czech IDS. Infants acquiring Czech are perceptually sensitive to vowel duration differences from an early age on: already at birth

their brains respond to a native vowel length contrast as strongly as to a native spectral contrast [11] and throughout the infants' first year of life, phonological length differences remain behaviourally discriminable at least as robustly as spectral contrasts [12]. Based on these developmental patterns, we hypothesize that the Czech infants' input facilitates the learning of vowel length, likely through the exaggeration of short-long differences in IDS. On the other hand, infants acquiring other languages with phonemic length master the native length contrasts even though studies have not detected exaggeration of length contrasts in their IDS [13]. An alternative hypothesis therefore is that Czech IDS will be slower overall, with similarly prolonged durations of both short and long vowels, and without an exaggeration of the duration-cued differences.

Regarding the developmental trajectory of infants' early input, one of the findings of [1] most relevant to our research question is that vowel duration in IDS appears to change over time: the older the infant gets, the shorter the vowels tend to be. It is worth pursuing this phenomenon further, considering the developmental changes associated with infants' perceptual acquisition of native vowel contrasts. As noted above, Czech infants' sensitivity to length contrasts seems to develop rather early, at or even before the 4th month of age. One could hypothesise that if durational contrasts are exaggerated in developmentally sensitive ways, larger exaggeration would be found for younger than for older infants. To address this, we investigate IDS spoken to infants across 4 to 10 months of age.

The phonological system of Czech contains ten monophthongal vowels, five short and five long phonemes. While duration is the primary cue for most Czech short-long contrasts, the high-front vowel contrast /i:/-/ɪ/ is cued also by spectral properties [14–16]. In some dialects, the spectral cue can even outweigh duration for /i:/-/ɪ/, and partially cues also the /u:/-/ʊ/ contrast [15, 17]. We therefore test whether the durational contrast in Czech IDS is realised differently in the high front vowels than in the other vowel pairs.

To sum up, we investigate whether (1) the durational difference cueing vowel length contrasts is exaggerated in Czech IDS, whether (2) the exaggeration depends on infant age, and whether it is (3) specific to vowel pairs for which vowel duration is the primary cue in ADS.

## 2. METHOD

### 2.1. Participants

Sixty-three mother-infant dyads participated in the experiment. The women were native speakers of Czech who had lived in the metropolitan Prague area for at least 5 years prior to the experiment. They were between 24 and 39 years old and had no speech or hearing disorders. Their children participating in the experiment were aged 4 ( $n = 18$ ), 6 ( $n = 13$ ), 8 ( $n = 19$ ), or 10 ( $n = 13$ ) months ( $\pm 2$  weeks).

### 2.2. Materials and procedure

Each of the 10 Czech monophthongs /ɪ i: ε ɛ: a a: o o: u u:/ was elicited in two words, altogether represented by a set of 20 objects which participants spontaneously commented on (following [16]). The target vowels always occurred in a word-initial syllable (which is, formally, stressed); the flanking consonants were bilabials and alveolars. Participants were instructed to name each object at least twice while talking about it. Participants talked about the set of objects twice, once to their infant (only the parent and the infant present in the recording booth) and once to an adult experimenter (with only the parent and the experimenter present), with order of the conditions counterbalanced. The recordings were done in a sound-treated booth using a condenser microphone AKG C520 and an Edirol UA 25 sound card, with Audacity run on a PC (44.1-kHz sampling frequency and 16-bit quantisation).

### 2.3. Acoustic analysis

Duration was measured over segmented word and vowel tokens. Only disyllabic words, and their corresponding first-syllable vowels, were included in the analysis. All tokens were manually segmented using Praat [18], based on visual inspection of the waveform and spectrogram. Word onset and offset corresponded to the onset of the first and the offset of the last segment. The target vowel interval had to include visible formants, particularly F2. All boundaries were aligned to zero crossing in the waveform. The final data set included 10249 target vowel tokens, 5782 of them in the IDS condition and 4467 in ADS roughly equally distributed across vowel categories.

## 2.4. Statistical analysis

Vowel durations were normalised for word duration, by dividing the duration of each vowel in milliseconds by the duration of the word in which it was embedded. The normalised vowel durations were submitted to a linear mixed-effects model using the *lme4* and *lmerTest* packages in R [19–21]. The modelled fixed effects were Style (ADS vs. IDS, coded as -1 vs. +1), Length (short vs long, coded as -1 vs +1), Vowel pair (with 4 sum-to-zero contrasts comparing the /i/-i:/ pair to each of the remaining 4 pairs), and Age (in months, mean-centered). The models contained a full random-effects structure with main and interaction effects of all the three within-subjects factors.

## 3. RESULTS

Table 1 lists the outcomes for the interaction effects involving Style and Length, since those can answer our research questions, as well as the simple main effects involved in those interactions.

Fixed effects	estimate	SE	t
<b>Intercept</b>	<b>0.193</b>	<b>0.0018</b>	<b>109.22</b>
<b>Style (-ADS +IDS)</b>	<b>0.009</b>	<b>0.0007</b>	<b>12.43</b>
<b>Length (-lo + sh)</b>	<b>0.054</b>	<b>0.0007</b>	<b>76.04</b>
<b>Age</b>	<b>-0.004</b>	<b>0.0024</b>	<b>-2.25</b>
<b>Style*Length</b>	<b>0.004</b>	<b>0.0070</b>	<b>5.38</b>
<b>Style*Length*Age</b>	<b>-0.002</b>	<b>0.0010</b>	<b>-2.04</b>
Sty*Len*V.pair(-a +i)	0.002	0.0014	1.35
Sty*Len*V.pair(-e +i)	-0.0005	0.0015	-0.31
Sty*Len*V.pair(-o+i)	-0.002	0.0014	-1.41
Sty*Len*V.pair(-u+i)	-0.0006	0.0015	-0.41

**Table 1:** Model outcome of selected fixed effects, namely, of interactions involving the factors Style and Length, and the related simple effects. DFs for Intercept and Age = 63, other DFs ~ 10200. Effects yielding  $p < 0.05$  are in bold.

The parameters relevant to our research question about realisation of the Czech length contrast in IDS are those that involve the interaction of Style and Length. The significant two way-interaction of Style and Length suggests that the length contrast is realised differently in IDS than in ADS:

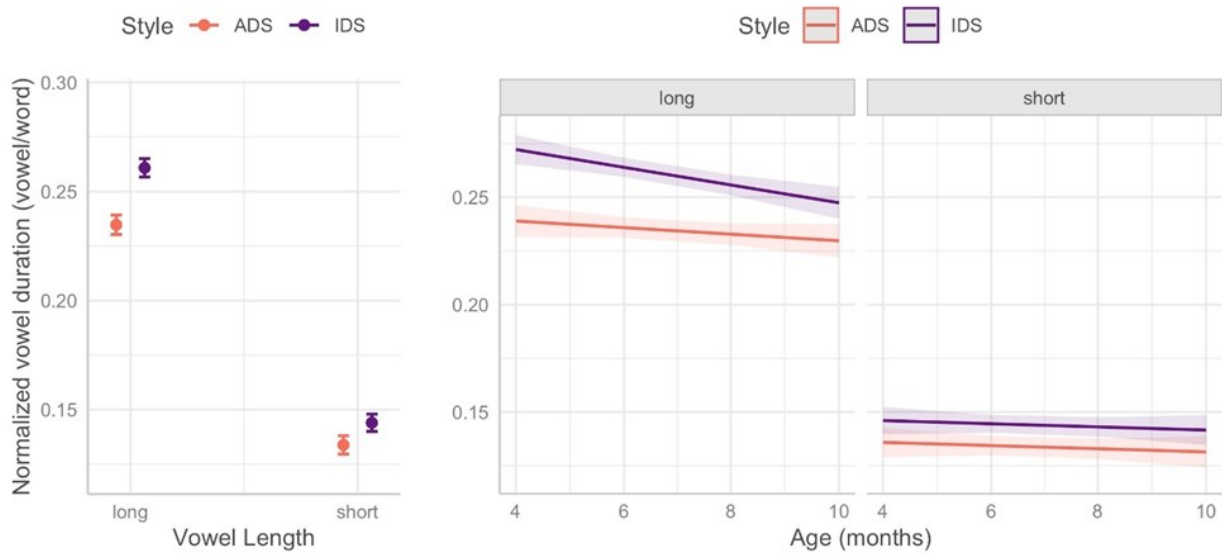
pairwise comparisons show that long vowels differ between IDS and ADS more than short vowels (see Fig. 1, left panel). This two-way interaction is further licenced by the significant three-way interaction additionally involving Age: as shown in Figure 1 (right panel), the greater lengthening of long vowels in IDS as compared to short vowels decreases with infant age.

None of the interactions of Style and Length with Vowel pair turned out significant. We thus failed to find evidence that the differential cue-weighting of duration for the high-front vowel pair documented for Czech adult speech would affect length contrasts in Czech IDS. Given this null result for the interaction of Style, Length, and Vowel pair we will not make any conclusions about vowel-quality-specific realisation of length in Czech IDS.

## 4. DISCUSSION

In this paper, we investigated whether in Czech, a quantity language, phonemic length contrasts are exaggerated in infant-directed speech. To this end, we measured the durations of phonemic short and long vowels in the speech of Czech mothers addressed to their 4- to 10-month-old infant and to an adult. The analyses revealed that not only are vowels longer in Czech IDS than in ADS in general (main effect of Style) but that the lengthening is larger for long vowel phonemes than for short phonemes (interaction of Style and Length) and increases with decreasing age of the infant who is being spoken to (three-way interaction of Style, Length, and Age). To our knowledge, such age-specific enhancement of length contrasts in IDS has not been widely reported for quantity languages, very probably because length contrasts as such have been investigated only in a few IDS studies to date [6, 8].

The present findings for Czech suggesting the exaggeration of length contrast in speech to very young infants (and less so to older ones) are partly in line with the results for exaggeration of vowel length distinctions reported earlier for Swedish IDS [6]. The developmentally conditioned early exaggeration of length distinctions in Czech, could have a facilitative function, aiding infants to acquire the contrast between short and long vowel phonemes. After all, infants acquiring Czech seem to be perceptually sensitive to native length contrasts at birth as well as at 4 months [11, 12], unlike for instance infants acquiring Japanese, who distinguish native vowel length contrasts



**Figure 1:** The estimated mean normalised vowel duration and its 95% confidence intervals for the two way-interaction of Style and Length (left) and for the three-way interaction of Style, Length, and Age (right).

much later, at 9.5 months [9].

A plausible explanation for this between-language difference in infants' perceptual abilities could lie in the input, where contrast exaggeration promotes early vowel length acquisition in Czech and low frequency of occurrence of long vowels postpones vowel length acquisition in Japanese [10]. It would be worth to further compare the realisation of length contrasts in IDS across various languages in which vowel length is phonemic, such as Slovak, Hungarian, Finnish, Arabic or Estonian, and monitor its relationship to the perceptual development of infants acquiring those languages.

Contrary to our predictions, according to which we expected the length contrast to be realised differently in IDS for the high-front vowel pair than for other pairs, we did not detect any such effects. This could mean that Czech speakers realise the length contrast in high front vowels similarly across ADS and IDS. But note that we did not analyse the participant's native dialects that might affect the realisation of this length contrast. In the present study, the mothers came from various parts of Czechia, the only criterion for inclusion in the study being that they had lived in the Prague area at least 5 years prior to the recording. At least for some vowel pairs, the weighting of vowel duration as a cue to vowel length contrasts varies between the western and eastern Czech dialects [14]. One could thus speculate that the realisation of durationally-cued contrasts may differ also in IDS across those dialects. In future work on IDS, dialectal variation should be taken into account.

The literature reports on cross-linguistically observed changes in IDS due to infant age, such as the overall shortening of vowel durations with increasing infant age, thus contributing to progressively faster speech tempo [1]. It remains to be investigated whether developmental changes in IDS also occur for spectral properties of vowels such as diphthongisation, for vowel space size, or for consonants, and to what extent they depend on the language, dialect, and infant age at hand.

## 5. CONCLUSION

When talking to their infants, Czech-speaking mothers not only produce longer vowels overall but also exaggerate the durational differences between phonologically short and phonologically long Czech vowels, as compared to when speaking to an adult. This exaggeration of length contrasts seems to be larger for speech addressed to younger infants than for speech addressed to older infants. Future research is needed to better understand how infant age or mother's native language variety modulate IDS.

## 6. ACKNOWLEDGMENT

This work was funded by Charles University grant Primus/17/HUM/19 and Czech Science Foundation grant 21-09797S. We are grateful to the participating families and also to Kristýna Hrdličková, Veronika Ungrová, and Zuzana Oceláková for help with data collection and annotation.

## 7. REFERENCES

- [1] Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., Fusaroli, R. 2022. A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, 1-20.
- [2] Chládková, K., Černá, M., Paillereau, N., Skarnitzl, R. and Oceláková, Z. 2019. Prenatal infant-directed speech: vowels and voice quality. *Proceedings of ICPHS 2019*, pp.1525-1529.
- [3] Benders, T. 2013. Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, 36(4), 847-862.
- [4] De Boer, B., Kuhl, P. K. 2003. Investigating the role of infant-directed speech with a computer model. *Acoustics Research Letters Online*, 4, 129
- [5] Cristia, A. 2013. Input to Language: The Phonetics and Perception of Infant-Directed Speech. *Language and Linguistics Compass* 7(3), 157–170
- [6] Sundberg, U. 1998. Mother tongue - Phonetic Aspects of Infant-Directed Speech (PhD dissertation, Department of Linguistics, University of Stockholm).
- [7] Englund, K., Behne, D. 2006. Changes in infant directed speech in the first six months. *Infant and Child Development: An International Journal of Research and Practice*, 15(2), 139-160.
- [8] Tajima, K., Tanaka, K., Martin, A., Mazuka, R. 2013. Is the vowel length contrast in Japanese exaggerated in infant-directed speech? *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 3211-321.
- [9] Mugitani, R., Pons, F., Werker, J.F., Amano, S. 2009. Perception of vowel length by Japanese and English-learning infants. *Developmental Psychology*, 45, 236– 247.
- [10] Bion, R. A., Miyazawa, K., Kikuchi, H., Mazuka, R. 2013. Learning Phonemic Vowel Length from Naturalistic Recordings of Japanese Infant-Directed Speech. *PLOS ONE* 8(2): e51594.
- [11] Chládková, K., Urbanec, J., Skálová, S., Kremláček, J. 2021. Newborns' neural processing of native vowels reveals directional asymmetries. *Developmental Cognitive Neuroscience*, 52, 101023.
- [12] Paillereau, N., Podlipský, V. J., Šimáčková, Š., Smolík, F., Oceláková, Z., Chládková, K. 2021. Perceptual sensitivity to vowel quality and vowel length in the first year of life. *JASA Express Letters*, 1(2), 025202.
- [13] Bernstein Ratner N, Luberoff A. 1984. Cues to post-vocalic voicing in mother-child speech. *Journal of Phonetics*, 12, 285–289.
- [14] Šimáčková, Š, Podlipský, V., Chládková, K. 2012. Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association*, 42(2), 225-232.
- [15] Podlipský, V. J., Chládková, K., Šimáčková, Š. 2019. Spectrum as a perceptual cue to vowel length in Czech, a quantity language. *The Journal of the Acoustical Society of America*, 146(4), EL352.
- [16] Paillereau, N. & Chládková, K. (2019). Spectral and temporal characteristics of Czech vowels in spontaneous speech. *AUC PHILOLOGICA*. 2019. 10.14712/24646830.2019.19.
- [17] Bořil T., Veroňková J. 2020. Perceived Length of Czech High Vowels in Relation to Formant Frequencies Evaluated by Automatic Speech Recognition. In: Sojka P., Kopeček I., Pala K., Horák A. (eds), *Text, Speech, and Dialogue. Lecture Notes in Computer Science*, 12284. Springer, Cham, 409-417
- [18] Boersma, P., Weenink, D. (2019). Praat: doing phonetics by computer (Version 6.0.25). <http://www.praat.org/>
- [19] Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- [20] Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software*, 82(13), 1–26.
- [21] R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.